



# RISEkbmRasch

an R package to simplify and integrate analysis and documentation with Quarto

Magnus Johansson, PhD   
RISE Research Institutes of Sweden

2023-06-15

[magnus.p.johansson@ri.se](mailto:magnus.p.johansson@ri.se)

# Agenda

- a brief background of the package
  - reproducibility, transparency & quality
- walkthrough of Quarto/R code and its output
  - using the package vignette as an example (available online)

**Useful links -  
[github.com/pgmj](https://github.com/pgmj)**

# Why R?

- open source & free
- potentially easier for others to reproduce and learn from your results
- available for Windows/Mac/Linux/etc

# But...

Plots and other output still needs copying to a text document to collect the output and document the analysis

- challenging to reproduce results
  - especially if you go back to an old project

# Quarto

<https://quarto.org>

- documentation in the same document as the analysis code!
- can output PDF, Word, HTML documents (and more)
- makes reproducibility simple
- (yes, it is like Rmarkdown)
- pre-installed with Rstudio

# Brief history

- I tried ordering RUMM but didn't get a response
- Took a course on IRT/CAT with R
- Created code to recreate/develop output of Winsteps/RUMM
- Workshop with colleagues - *"Maybe make functions to simplify?"*

# Before & after functions

## Before

```
df.erm<-PCM(df.omit.na) # run PCM model, replace with RSM (rating scale) or
RM (dichotomous) for other models
# get estimates, code borrowed from
https://bookdown.org/chua/new\_rasch\_demo2/PC-model.html
item.estimates <- eRm::thresholds(df.erm)
item_difficulty <- item.estimates[["threshtable"]][["1"]]
item_difficulty<-as.data.frame(item_difficulty)
item.se <- item.estimates$se.thresh
person.locations.estimate <- person.parameter(df.erm)
item.fit <- eRm::itemfit(person.locations.estimate)
std.resids <- item.fit$st.res
# PCA of Rasch residuals
pca <- pca(std.resids, nfactors = ncol(df.omit.na), rotate = "oblimin")
# create table with top 5 eigenvalues
pca$values %>%
  round(2) %>%
  head(5) %>%
  as_tibble() %>%
  rename('Eigenvalues' = 'value') %>%
  kbl(booktabs = T, escape = F, table.attr = "style='width:25%;'" ) %>%
# options for HTML output
kable_styling(bootstrap_options = c("striped", "hover"),
              position = "center",
              full_width = T,
              font_size = r.fontsize,
              fixed_thead = F) %>%
column_spec(1, bold = T) %>%
kable_classic(html_font = "Lato") %>%
# latex_options are for PDF output
kable_styling(latex_options = c("striped","scale_down"))
```

## After

```
1 RIpcmPCA(df)
```





# What is an R package?

- a collection of `functions()`
- `RISEkbmRasch` relies 100% on other packages
  - it can be described as a “wrapper” package
  - it is also an “opinionated” package

# Package ambitions

- make it as simple as possible to get key tables and figures
- you can choose cut-off values for highlighting in most functions, for instance:
  - item fit over/under a certain value
  - residual correlations relative to average residual correlations
- more flexibility gradually added (but also adds complexity)

# Notes on choices

There are multiple R packages for Rasch analysis.

- We went with **eRm** primarily
  - handles dichotomous and polytomous data
  - uses CML, conditional maximum likelihood
    - “specific objectivity”
    - ordinal sum score as a “sufficient metric”
- **mirt** for Yen’s Q3 residuals
- **psychotree** for DIF (differential item functioning)

# Simulation study coming

Partial credit model (PCM) analysis

- **eRm** with CML
- **TAM** with MML

Comparisons with variation in sample size and targeting.  
May also produce a reasonable basis for assessing power for Rasch analysis. R code will be included.

# What's in the package?

- Descriptive analysis
  - distribution of data
  - missing data
  - Guttman “heatmap”

# Required data structure

- one dataframe with item data **ONLY**
  - coded as integers starting with 0 for lowest response category
- one dataframe with item descriptions
- (DIF variables as separate vectors)

# Rasch-related functions

- Let's go to Rstudio and the sample code from the package vignette
  - all code is available on GitHub



# Benefits

- you can make a template analysis file
  - makes it harder to miss important steps
    - quality assurance
- easier for others to understand your analysis process, step by step
  - easier for yourself to go back to old analyses...
- transparency in decision making
- complete reproducibility if data is shared

# Report everything!?

- You can share a fully documented report file as an appendix document with the preprint
- Example:

Rozental, A., Forsström, D., & Johansson, M. (2023). A Psychometric Evaluation of the Swedish Translation of the Perceived Stress Scale: A Rasch Analysis [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-2699284/v1>

# A note on templates

Our group at RISE have made an analysis template based on our preprint, in which we propose a reporting standard for psychometric analyses. It builds on Tennant & Conaghan's 2007 paper and others.

Johansson, M., Preuter, M., Karlsson, S., Möllerberg, M.-L., Svensson, H., & Melin, J. (2023). *Valid and Reliable? Basic and Expanded Recommendations for Psychometric Reporting and Quality Assessment*. OSF Preprints.

<https://doi.org/10.31219/osf.io/3htzc>

